

# Spectral estimation of HRV in signals with gaps

L. Rodríguez-Liñares<sup>a,\*</sup>, D.M. Simpson<sup>b</sup>

<sup>a</sup>*Department of Computer Science, Universidade de Vigo, 32004 Ourense, Spain*

<sup>b</sup>*Institute of Vibration and Sound Research, University of Southampton, Southampton SO17 1BJ, United Kingdom*

---

## Abstract

Heart rate variability is commonly quantified following spectral estimation. However, it is often difficult to obtain continuous recordings of beat-to-beat intervals without interruptions due to artefacts, noise or sporadic arrhythmias. Such data loss may be seen as gaps in the recordings, and often results in such signals being discarded. While a number of methods has been proposed for spectral estimation in heart rate records with gaps, there are no comprehensive comparisons between them. This paper tries to fill this void, comparing methods and identifying the most versatile and reliable one.

The mean (bias error) and standard deviation (random error) of estimates of power in the low frequency band (LF), from 0.04 to 0.15 Hz; in the high frequency band (HF), from 0.15 to 0.4 Hz; and their ratio (LF/HF), were calculated in RR-interval time-series with up to 50% of samples missing through large or small gaps introduced into recordings.

‘Correlogram (bridging)’ and ‘Burg for segments’ methods proved to be the most robust methods for dealing with gaps, but Burg for segments was found to be more robust, especially in the HF band. Our results clearly show that even large gaps (covering a total of 50% of the recording time) can still yield robust spectral estimates of HRV, provided appropriate methods are used.

*Keywords:* Heart rate variability, biomedical signal processing, electrocardiography, spectral analysis, data loss

---

## 1. Introduction

Cardiovascular diseases are the most important cause of death in the world [1]. Consequently, much effort has been devoted to providing systems and algorithms to help to detect and classify cardiac disease and dysfunction of cardiovascular control, including that of the autonomic nervous system [2].

---

\*Corresponding author

*Email addresses:* leandro@uvigo.es (L. Rodríguez-Liñares), ds@isvr.soton.ac.uk (D.M. Simpson)

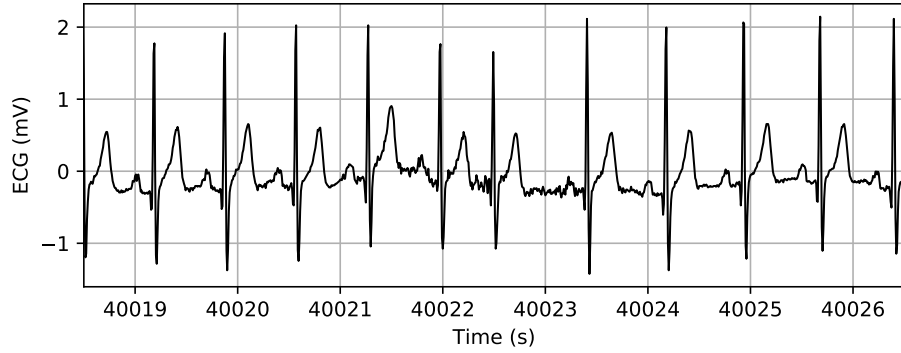
Heart rate variability has become a much used tool in this area [3, 4] as well as in others, such as psychophysiology [5].

Heart rate (HR) is calculated as the number of heart-beats per unit of time, usually per minute. However, heart rate is not constant, presenting variations caused by a range of environmental, physiological and psychological challenges, which stimulate the various cardiovascular control mechanisms [6]. Heart rate variability (HRV) refers to these beat-to-beat alterations in heart rate (or more commonly in the time interval between successive heart beats). It is most commonly measured from the Electrocardiographic signal (ECG), based on the interval between successive R-waves (RR interval). For the assessment of autonomic nervous system function, spectral analysis is then often carried out. In such analyses, LF (low frequency) components (from 0.04 to 0.15 Hz), are commonly assumed to be influenced by both sympathetic and parasympathetic systems, whereas HF (high frequency) components, (from 0.15 to 0.4 Hz), estimate the parasympathetic tone and are primarily related to the respiratory rhythm [6]. The LF/HF ratio can then be used as an index of the sympathovagal balance [4].

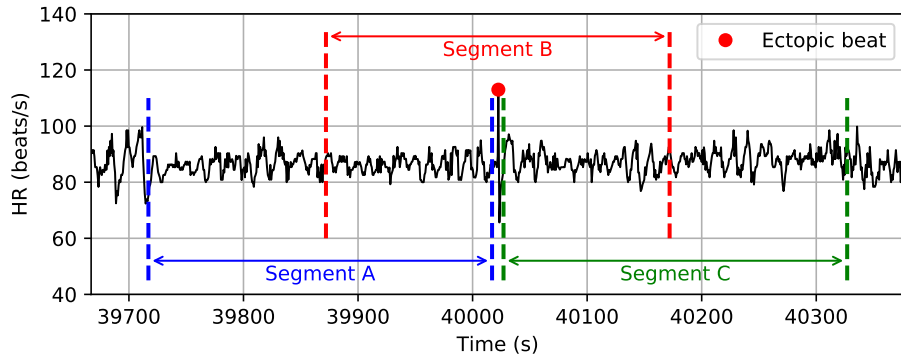
A major problem in the analysis of HRV (as well as in many other biomedical signals) are the often unavoidable artefacts, which can render the data useless. In the case of the ECG, which is the most common signal from which HRV is found, the artefacts can be provoked by external sources, such as electrical interferences, or internal sources, such as muscle activity or patients' movements. In addition, some arrhythmias, including occasional ectopic beats, represent an important source of distortion in spectral estimates that may occur in both normal subjects and patients with heart disease and should be generally excluded from spectral estimation [6, 7].

The importance of even one aberrant beat in spectral estimation was highlighted in [8]. To illustrate this point, Fig. 1 shows the spectral distortion introduced in a heart-beat record with just one ectopic beat, characterized by an unusually short RR interval followed by a long one. We selected a record from the MIT-BIH Normal Sinus Rhythm Database that contains four such ectopic beats and used a procedure similar to the one used in [7] to assess their impact: we compared the AR (autoregressive) spectra of a segment containing the ectopic beat (segment B), against the previous (segment A) and the following (segment C) segments. Each segment is approximately 5 minutes in duration, as recommended [4]. Given that the subject was at rest, similar spectra might be expected, and that is indeed observed for segments A and C, but segment B shows large deviations in the region from about 0.2 Hz upwards. Results such as these (and worse cases) motivate the current paper.

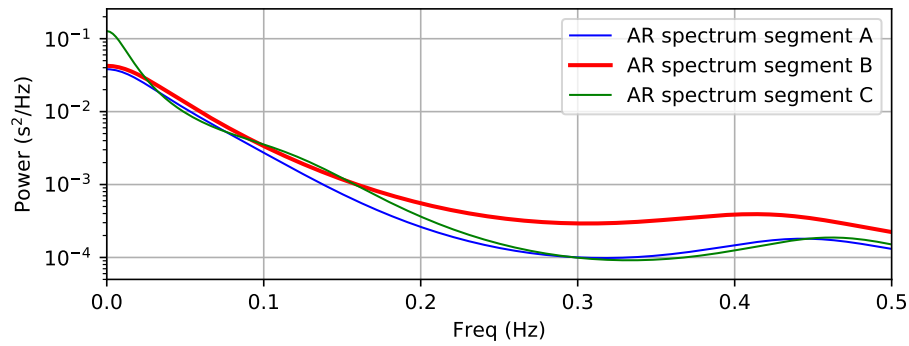
Some techniques have been proposed in the literature to deal with ectopic beats such as interpolation [9], integral pulse frequency modulation (IPFM) [7], phantom beats [10] or non-linear filtering [11]. However, the main drawback of many of these approaches is that they do not deal adequately with longer gaps in the normal heart-rhythms that may be introduced by artefacts, data loss or changes in physiological status, e.g. due to intermittent changes in respiratory or activity patterns.



(a)



(b)



(c)

Figure 1: Record 16483 of the MIT-BIH Normal Sinus Rhythm Database: (a) ECG segment with ectopic beat, (b) heart rate of the segment containing the ectopic beat, and (c) AR-spectra of segments before, after, and centered on the ectopic beat.

The current paper brings together and compares a number of different methods for spectral estimation with gapped data, for the specific application on HRV time-series. The objective is to assess methods when up to 50% of the data is missing. Such records with large amounts of missing data are often discarded, or shortened to using only the longest good quality segments. Alternatively, short gaps are commonly interpolated from adjacent samples, followed by the application of standard spectral estimation algorithms. These approaches are either inefficient in terms of data usage or raise some concern about the robustness of results [12]. This paper is focused on testing established methods such as Burg for segments [13] and Lomb-Scargle [14], and in addition some modifications of standard spectral estimation approaches. The aim is to assess if spectral estimators for signals with gaps improve on these more commonly used alternatives and which of the methods should be recommended.

## 2. Methods

When trying to perform spectral estimation from missing data, two alternatives arise: a) reconstruct the data by estimating (imputing) samples in the gaps and then proceed as if the signal had no gaps, or b) use a spectral estimation method that uses only the available data and can deal with interrupted (gapped) recordings. The first group includes methods such as concatenation or variants of interpolation (nearest neighbour, linear, cubic, splines, etc.) [9, 12, 15]. These methods are often used, but their acceptability depends on the characteristics of the data and the size and distribution of the gaps present in the signal [12]. The second group estimates the spectral characteristics of the signal by using variants of spectral estimators tailored to deal with the gaps. These include the correlogram and Yule-Walker [16], Burg algorithm [17], as well as the Lomb-Scargle method [14, 18], where in the latter the gaps are viewed as a special case of irregular sampling. The Welch method, which is probably the most used technique for HRV spectral estimation, cannot be applied directly in data with gaps because the discrete Fourier transform (DFT) requires all samples in each window; thus it is not included in the analysis.

### 2.1. Concatenation

The simplest approach when there are gaps or unusable portions in the signal is to ignore them and join the remaining segments [9, 15]. This procedure evidently introduces large transients in the time-domain signal and hence spectral distortion, as it will be shown in the Results section.

Several studies have previously recommended avoiding this method [15, 19, 20], but is included here to assess the impact of this approach, whose evident benefit is simplicity and thus may be tempting.

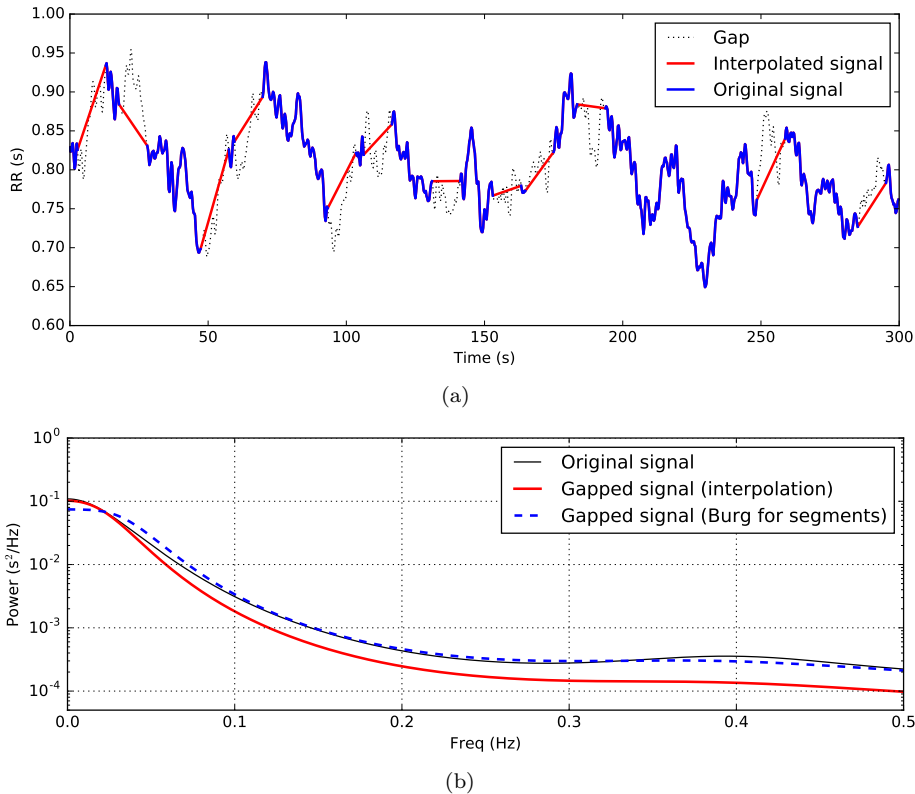


Figure 2: (a) An example of a complete RR interval signal and the same with gaps (grey) replaced by linear interpolation (red). (b) The corresponding spectra of the complete signal (black), of the interpolated version (red) and the approach recommended in this paper that does not require interpolation (Burg for segments – see below, in blue).

## 2.2. Interpolation

Whereas concatenation effectively ignores the gaps and joins the remaining segments, interpolation attempts to fill the gaps with estimated values.

This approach can be useful when dealing with short gaps such as caused by ectopic beats [9, 15, 21], but it is obvious that this is not suitable for long ones. For comparison purposes, we used linear interpolation in our tests. Fig. 2 shows an interpolated RR sequence where gaps that are 10 s long covering a total of 40% of the signal.

One problem when using linear interpolation is that it introduces discontinuities at both ends of the gaps. Higher order interpolation methods attempt to alleviate this problem, replacing missing data by smooth segments ensuring continuity with the remaining signal. However, our tests with cubic interpolation showed that, as commented in [22], it can also result in huge excursions (over- or under-shoots) in the added segments that invalidate using this method, even for some relatively short gaps.

### 2.3. Correlogram

The correlogram method of spectral estimation is a well-known method that first estimates the autocorrelation sequence of the signal and then uses the Fourier transform to obtain the power spectral density (PSD) [16]. In practice, this means finding the product of samples with a fixed lag and averaging these over the full duration of the signal and then repeating this process over lags from zero to a maximum value chosen by the user. This may be viewed as an operation with a sliding window (whose length corresponds to the chosen maximum lag in the autocorrelation function) passing over the signals, and for each sample in the signal, its product with all samples in the window is calculated. The final autocorrelation is given by the average of these products for each lag.

Gaps are easily introduced in this method as zeros, allowing conventional methods for cross-correlation estimate to be applied. The effect of the zeros must then be compensated for by correcting the denominator in averaging, excluding all samples that involve the gaps (zeros). However, it may be possible that some frames (windows) enclose a gap in its entirety with non-zero signal values at each side of the gap. How to deal with these cases gives two possible variants of this method:

(i) *Bridging over the gaps:*

In this case, the correlogram method is applied without modification: some frames will include gaps and portions of valid signal on both sides. As a consequence, the number of samples used at each lag may vary erratically, as some samples may be used at small and large lags, but not for intermediate values.

(ii) *Not bridging over the gaps:*

Gaps are expanded by inserting zeros to ensure that the signal segments at both sides of gaps are always in different frames. As a result, the number of samples used at each lag decreases monotonically, as it would in signals without gaps. Once a sample is excluded at one lag in the current window (because the lagged sample falls into a gap), it cannot be included again at a larger lag, even if in the original recording that sample was available, at the other side of the gap. This approach is thus expected to lead to less erratic variations in the autocorrelation estimate than the previous method. One disadvantage of this approach is that some information is discarded, with the correlation between samples before and after gaps not being exploited.

In the current work, the correlogram method was applied by sliding a boxcar window of length 120 samples over the signals, i.e. the autocorrelation function was limited to  $\pm 120$  samples.

### 2.4. Yule-Walker method

Yule-Walker AR estimation method (or autocorrelation method) tries to fit an autoregressive model by sliding a window over the input data and minimizing

the forward prediction error [16]. This leads to the Yule-Walker equations, which can be solved efficiently using Levinson’s recursion. This may be viewed as closely related to the correlogram method, but rather than assuming that the autocorrelation is zero for lags greater than the window (frame) size, the autocorrelation is now extrapolated based on an autoregressive model fitted to the first few lags (corresponding to the model order) of this function. In both these methods the signals are being viewed as the result of filtering white noise. In the case of the correlogram method, this filter is a finite impulse response (FIR) filter, for Yule-Walker, this is an autoregressive (infinite impulse response – IIR) filter.

As with the correlogram method, there are again two variants of this approach, which differ in the way they deal with data around the edges of gaps: taking them into account for the estimation (bridging the gaps) or not (not bridging the gaps).

### 2.5. Lomb-Scargle method

Lomb proposed a method for estimating Fourier transforms of unequally spaced segments of signals [14]. This method was later improved by Scargle, who applied it to the problem of detection of signals in astronomical time series [18]. Signals with gaps may be considered signals where observation times are not evenly spaced.

The key idea of the Lomb–Scargle method is to try to estimate an individual Fourier coefficient for every frequency of interest. For each frequency, it computes the spectral density from the amplitudes of a least-squares fit of the sum of a sine and a cosine fitted to the available observations. This works well when the signal under study is composed of a periodic component contaminated by noise, but the method is limited in modelling the shape of the spectrum [10, 19, 23].

### 2.6. Burg algorithm for segments

In AR modelling, the Burg algorithm [17] is often regarded as the preferred estimator. It involves the direct estimation of so called reflection coefficients  $k_p$  that minimize the sum of the forward and backward prediction errors, using a recursive algorithm.

The Burg algorithm has been adapted to be used in cases where data is not continuous but given in separate segments. The usual approach in these cases is to take the average of the coefficients of the models estimated from the individual segments [24]. This approach tends to introduce bias when the number of segments increases. For this reason, we used the solution proposed in [13], that consists of estimating the reflection coefficients  $k_p$  by minimizing the fit of the  $AR(p)$  model to all segments simultaneously. As the correlogram without bridging and the Yule-Walker method, this approach does not take the correlation between samples before and after gaps into account.

Record	Length	Beats	Segment
16265	22:14:22	100243	38000 - 38300
16483	21:08:19	104334	46500 - 46800
17052	21:13:20	87356	24000 - 24300
18177	24:15:30	115911	28000 - 28300
19088	20:10:13	97961	49500 - 49800
19830	21:06:05	109332	29000 - 29300

Table 1: Segments from the MIT-BIH Normal Sinus Rhythm Database selected to act as seeds to create the surrogates.

### 3. Experimental setup

In order to test and compare all the algorithms, we used a Monte-Carlo approach: we computed AR parameters from HRV records available in a public database [25] and from these we generated sets of surrogate data. In this data we introduced gaps to test the capability of the various algorithms to still provide robust estimates. The great benefit of surrogate data analysis is that the ‘true’ spectrum is known in advance, and all estimates can be compared to this. This overcomes the problem of using recorded HRV signals where there is always doubt as to which of the estimates is ‘correct’ or ‘better’. Surrogate data analysis also allows the statistics of estimation errors (bias, random errors) to be assessed from a large ensemble of signals with identical statistical properties.

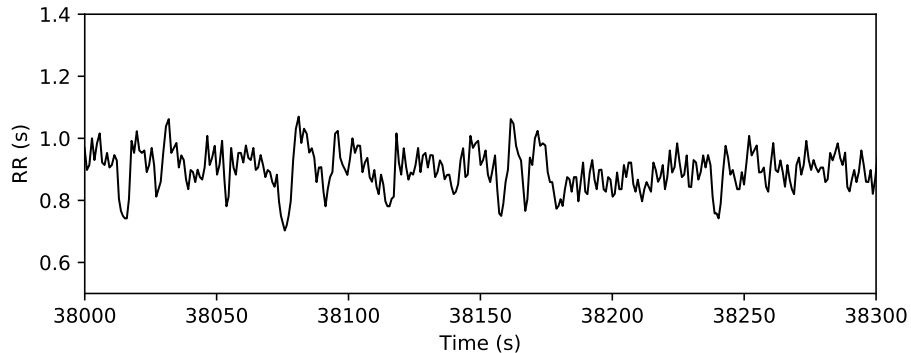
The work was carried out with in-house software, using the Python programming language and Spectrum libraries [26]. For the Lomb-Scargle estimate, we used the WFDB software package [27] available from the Physionet repository [25].

#### 3.1. Creating the surrogates

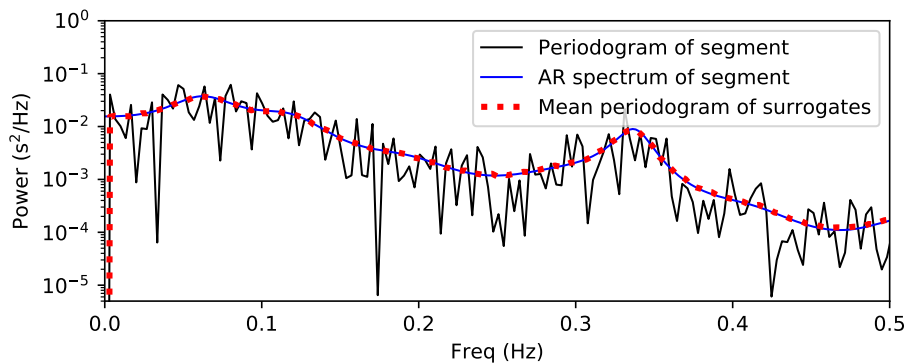
The MIT-BIH Normal Sinus Rhythm Database [25] includes annotated long-term recordings of subjects with no significant arrhythmias. From this database, we selected six recordings and the RR annotations provided with them. We selected one segment from each recording following the recommendations for HRV spectral analysis given in [4]: 300 s (5 minute) segments obtained under physiologically stable conditions (see Table 1). The segments were randomly selected after visual inspection to guarantee that they met the criteria of being physiologically stable and without artefacts. The selection was made blindly, with no a priori knowledge of their spectral characteristics.

Fig. 3 shows the segment extracted from recording 16265. A different set of segments could have been tried, but considering that their role is only to act as seeds for the surrogate data, similar results can be expected from a different selection of ‘typical’ recordings. The segments were linearly interpolated at 4 Hz and their mean values removed, such that only the variability in heart rate is analyzed. Note that the Lomb-Scargle method was also applied to the interpolated signals (with or without gaps), and not simply to the sequences of





(a)



(b)

Figure 3: Segment from recording 16265: (a) original signal, and (b) spectra of original segment (periodogram and AR order 50) vs mean periodogram of surrogates.

RR intervals for each heart-beat record (as used in some previous works [28]), to facilitate comparison with other methods.

From each of these segments, we estimated AR parameters using the Burg method. Then, using each of the AR models and independent white Gaussian noise, we generated 1000 surrogate signals of length 300 s (1200 samples) thus giving a total of 6000 surrogate records. AR model orders between 8 and 20 have been recommended for HRV signals [4, 29, 30] and for the current data the Akaike Information Criterion (AIC) suggested model orders greater than 10. However, in order to avoid bias in the comparison of spectral estimates which may arise from generating AR signals with the same model order as will be used in the AR spectral estimates, we opted to simulate the signals with an exaggerated model order of 50. This will give spectra with more detail than can be captured by the lower order AR spectral estimates (see Results section) and allow a fairer comparison between techniques. Fig. 3 illustrates the simulations by comparing periodograms and AR spectra obtained from one original segment

(recording 16265) and its surrogates.

### 3.2. Introducing gaps

Many methods to deal with gaps are focused on short interruptions in the signals (like gaps provoked by ectopic beats). In our research, we wanted to test also the robustness of the methods studied with long gaps typically produced by artefacts, such as those due to movement, muscle activity, poor contact or other sporadic loss of data.

For this purpose, we used two gap lengths: 3 and 10 seconds. Average RR intervals in the selected segments range from 0.6 to 0.9 s, approximately. Introducing 3 s gaps assures that, in all cases, the removed segments include at least a beat and the two adjacent ones, as it is required to remove completely the spectral distortion caused by an ectopic beat. For larger gaps, we assumed 10 s to be a reasonable value to simulate sporadic loss of data.

In both cases, we generated gaps in the surrogates covering 10, 20, 30, 40 and 50% of the signal duration. The gaps were inserted into the data iteratively: a position in the file was randomly selected and a gap inserted if it did not overlap with a previous one; if there was overlap, another position was selected. As the interpolation frequency was 4 Hz, 3 s and 10 s gaps covered 12 and 40 samples, respectively.

### 3.3. Assessing the algorithms

Spectral estimation in short-term HRV recordings are commonly analyzed according to three main spectral bands: VLF (Very Low Frequency, 0.003 to 0.04 Hz), LF (Low Frequency, 0.04 to 0.15 Hz), and HF (High Frequency, 0.15 to 0.4 Hz) [4]. Among these bands, the most used in HRV studies are the higher ones (LF and HF), which were included in our evaluation of the spectral estimation methods, along with the LF/HF ratio.

For each estimation method, each subject and each gap distribution pattern, mean and standard deviation of power across surrogates were calculated for bands LF and HF and for the LF/HF ratio.

In order to compare numerically the different methods, we needed a compact way of assessing the algorithms. For this reason, we used the mean Normalized absolute Bias Error ( $\overline{|NBE|}$ ) and mean Coefficient of Variation ( $\overline{CV}$ ) for each estimation method, each gap pattern and each band. These parameters were calculated as follows:

- (i) *Mean Normalized absolute Bias Error*

$$\overline{|NBE|} = \text{mean}_{\text{surrogates}} \left| \frac{S_{\text{gaps}} - S}{S} \right|$$

where  $S_{\text{gaps}}$  is the estimate of the parameter under study (LF and HF power and their ratio) obtained from the surrogate with gaps and  $S$  is the same parameter obtained from the signal without gaps. The absolute values of NBE are used, since the difference ( $S_{\text{gaps}} - S$ ) can be positive or

negative, which would lead to a smaller mean value. The mean value was calculated by averaging over all 6000 surrogates.

(ii) *Mean Coefficient of Variation*

$$\overline{CV} = \text{mean}_{\text{subjects}} \left| \frac{sd(S_{gaps})}{\text{mean}(S_{gaps})} \right|$$

where the mean and standard deviation (sd) of  $S_{gaps}$  are obtained from the 1000 surrogates belonging to each subject, and their ratio is averaged across subjects.

$\overline{|NBE|}$  quantifies the bias error induced by the gaps and its value corresponds to an error/signal ratio, whereas  $\overline{CV}$  quantifies the random scatter in estimates.

## 4. Results

Figs. 4 and 5 show the mean and standard deviation of power in bands LF and HF and the LF/HF ratio obtained with all the methods for surrogates for the subject 16265 (the other subjects gave similar results) when the percentage of gaps increases. Fig. 4 corresponds to 3 s gaps, whereas Fig. 5 includes results for 10 s gaps (i.e. less frequent but longer gaps). Also included is the value obtained from estimating PSD using the periodogram method on the surrogates with no gaps. The left plots represent bias error whereas the right plots show the random error in the estimates, and as expected, there is a general increase in standard deviation of estimates as the amount of data lost increases. There are some cases where the standard deviation decreases with increasing data loss, but this generally occurs when the mean value of the estimates also drops. Very clearly, some of the methods are better than others in providing a stable estimate with increasing amount and number of gaps. Note that results could be quite erratic for the correlogram not bridging over gaps and Yule Walker bridging over gaps, leading to very high standard deviations. For the sake of clarity these results were therefore not included in the plot.

Table 2 summarizes the results averaged across all surrogates for each gap pattern. It also includes, for each gap length, the average results across the different percentages of gaps (from 10 to 50%). Each cell contains two values:  $\overline{|NBE|}$  (up) and  $\overline{CV}$  (bottom). Best results of these parameters for each gap pattern (and their average) are emphasized by shading. It should be noted that the bias for each method at 0% gaps is zero, since bias is always calculated with respect to this condition.

### 4.1. Concatenation and linear interpolation

For these approaches, the spectral power was calculated using the periodogram method.

As said before, concatenation is the method used when ectopic beats are simply removed from the signal. This is disastrous for spectral estimation:

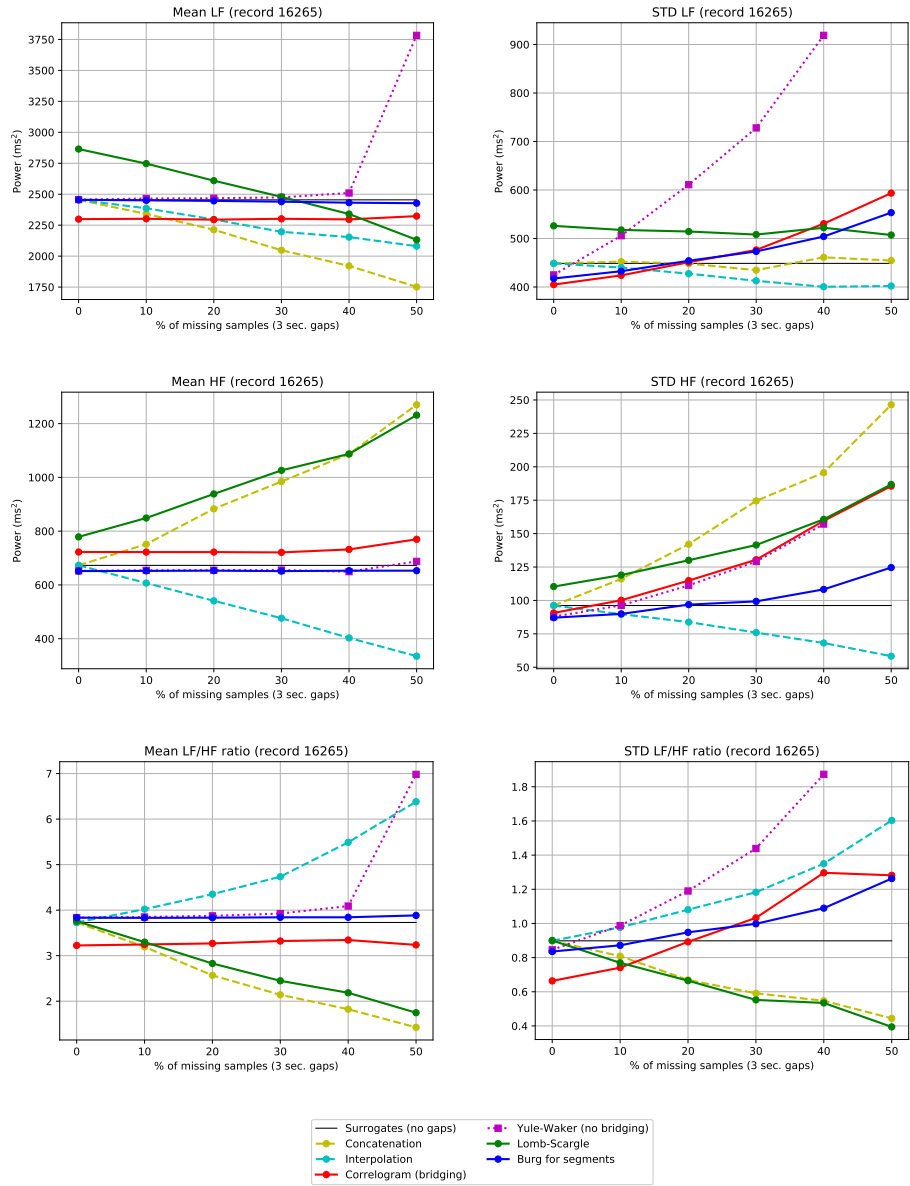


Figure 4: Mean and standard deviation of power for LF and HF bands and LF/HF ratio for each method for surrogates of record 16265 (3 s gaps).

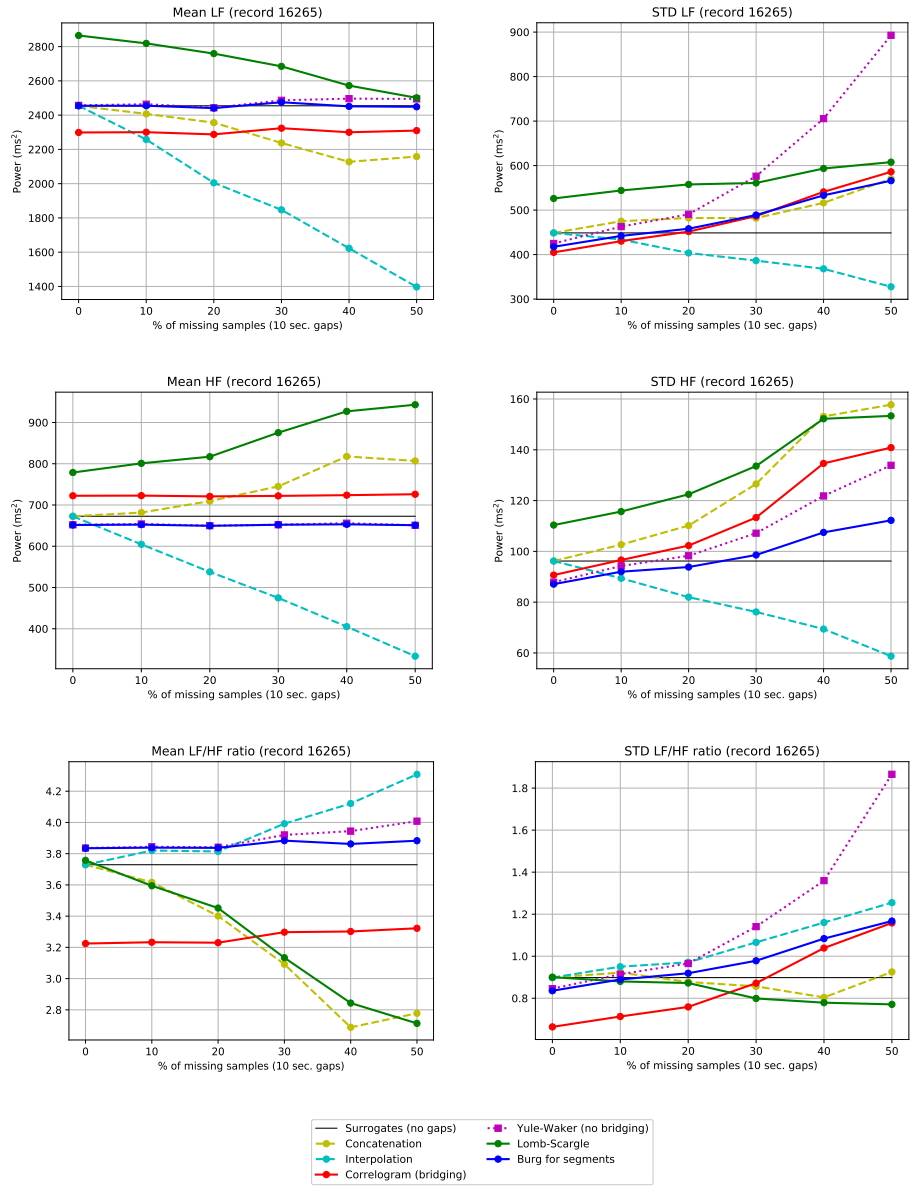


Figure 5: Mean and standard deviation of power for LF and HF bands and LF/HF ratio for each method for surrogates of record 16265 (10 s gaps).

		No gaps	Missing data (3 s gaps)					Missing data (10 s gaps)							
			10%	20%	30%	40%	50%	Avg.	10%	20%	30%	40%	50%	Avg.	
LF	Concatenation	0.000 0.198	0.076 0.208	0.117 0.215	0.167 0.232	0.211 0.250	0.320 0.272	0.178 <i>0.236</i>	0.081 0.208	0.107 0.223	0.144 0.237	0.176 0.252	0.255 0.278	0.153 <i>0.240</i>	
	Linear interpolation	0.000 0.198	0.061 0.201	0.127 0.203	0.180 <b>0.206</b>	0.219 0.211	0.281 0.213	0.174 <i>0.207</i>	0.084 0.206	0.168 0.216	0.234 0.228	0.318 0.244	0.387 0.255	0.238 <i>0.230</i>	
	Correlogram (bridging)	0.000 0.174	0.001 0.187	0.004 <b>0.202</b>	0.005 0.220	0.008 0.245	0.018 0.275	0.007 <i>0.226</i>	0.001 0.185	0.003 0.199	0.006 0.227	0.006 0.255	0.008 0.277	0.005 <i>0.229</i>	
	Correlogram (no bridging)	0.000 0.174	0.001 0.195	0.005 0.230	0.078 0.280	0.112 0.317	0.590 0.517	0.157 <i>0.308</i>	0.001 0.185	0.003 0.198	0.006 0.222	0.005 0.250	0.033 0.275	0.010 <i>0.226</i>	
	Yule-Walker (no bridging)	0.000 0.192	0.002 0.215	0.006 0.255	0.011 0.309	0.013 0.368	0.159 1.812	0.038 <i>0.592</i>	0.002 0.205	0.004 0.221	0.004 0.248	0.007 0.281	0.007 0.328	0.005 <i>0.257</i>	
	Yule-Walker (bridging)	0.000 0.192	0.100 1.458	0.068 0.643	0.226 2.723	0.515 4.457	0.631 4.830	0.308 <i>2.822</i>	0.167 2.267	0.068 0.611	0.144 1.624	0.317 3.793	0.404 4.416	0.220 <i>2.542</i>	
	Lomb-Scargle	0.000 0.194	0.166 0.195	0.338 <b>0.202</b>	0.537 0.210	0.783 0.225	1.095 0.241	0.584 <i>0.215</i>	0.142 0.201	0.308 0.210	0.592 0.221	0.915 0.233	1.307 0.248	0.653 <i>0.223</i>	
	Burg for segments	0.000 0.190	0.001 0.201	0.003 0.213	0.006 0.231	0.007 0.254	0.007 0.304	0.005 <i>0.241</i>	0.002 0.200	0.004 0.211	0.004 0.224	0.002 0.244	0.006 0.265	0.004 <i>0.229</i>	
	HF	Concatenation	0.000 0.159	0.159 0.173	0.416 0.190	0.633 0.205	0.846 0.215	1.221 0.226	0.655 <i>0.202</i>	0.078 0.175	0.134 0.186	0.228 0.196	0.395 0.221	0.429 0.231	0.253 <i>0.202</i>
		Linear interpolation	0.000 0.159	0.086 0.163	0.174 0.169	0.265 0.174	0.360 <b>0.181</b>	0.451 0.188	0.267 <i>0.175</i>	0.096 0.164	0.189 0.172	0.277 0.176	0.372 0.188	0.476 0.199	0.282 <i>0.180</i>
Correlogram (bridging)		0.000 0.137	0.002 0.173	0.012 0.230	0.056 0.269	0.100 0.293	0.426 0.337	0.119 <i>0.260</i>	0.002 0.152	0.001 0.163	0.002 0.189	0.008 0.236	0.028 0.250	0.008 <i>0.198</i>	
Correlogram (no bridging)		0.000 0.137	0.002 0.164	0.006 0.209	0.230 0.301	0.387 0.348	1.958 0.582	0.517 <i>0.321</i>	0.002 0.151	0.001 0.162	0.002 0.185	0.005 0.231	0.055 0.243	0.013 <i>0.194</i>	
Yule-Walker (no bridging)		0.000 0.139	0.003 0.156	0.007 0.182	0.015 0.234	0.020 0.296	0.134 0.817	0.036 <i>0.337</i>	0.002 0.149	0.003 0.160	0.004 0.177	0.007 0.199	0.012 0.232	0.005 <i>0.183</i>	
Yule-Walker (bridging)		0.000 0.139	0.239 3.529	0.070 1.686	0.065 1.374	0.189 2.185	0.579 3.283	0.228 <i>2.411</i>	0.023 0.769	0.030 0.553	0.041 1.164	0.089 1.755	0.222 3.741	0.081 <i>1.597</i>	
Lomb-Scargle		0.000 0.153	0.390 0.174	1.016 0.190	1.914 0.205	2.746 0.205	5.011 0.218	2.215 <i>0.198</i>	0.202 0.167	0.415 0.173	0.877 0.180	1.532 0.201	2.160 0.205	1.037 <i>0.185</i>	
Burg for segments		0.000 0.139	0.001 0.146	0.002 0.157	0.002 0.169	0.003 0.182	0.009 0.216	0.004 <i>0.174</i>	0.001 0.146	0.002 0.153	0.002 0.159	0.002 0.173	0.003 0.188	0.002 <i>0.164</i>	
LF/HF		Concatenation	0.000 0.255	0.162 0.266	0.307 0.276	0.404 0.295	0.485 0.318	0.594 0.341	0.391 <i>0.299</i>	0.117 0.274	0.151 0.280	0.222 0.304	0.301 0.320	0.316 0.356	0.221 <i>0.307</i>
		Linear interpolation	0.000 0.255	0.113 0.252	0.253 0.252	0.413 <b>0.251</b>	0.639 0.251	1.012 0.254	0.486 <i>0.252</i>	0.089 0.260	0.121 0.270	0.203 0.276	0.302 0.288	0.349 0.305	0.213 <i>0.280</i>
	Correlogram (bridging)	0.000 0.198	0.013 0.246	0.027 0.319	0.049 0.374	0.052 0.416	0.184 0.448	0.065 <i>0.360</i>	0.003 0.220	0.007 0.236	0.020 0.284	0.033 0.347	0.028 0.388	0.018 <i>0.295</i>	
	Correlogram (no bridging)	0.000 0.198	0.009 <b>0.242</b>	0.017 0.312	0.086 0.373	0.144 0.435	0.435 0.472	0.138 <i>0.367</i>	0.003 0.220	0.006 <b>0.235</b>	0.018 0.278	0.033 0.343	0.021 0.376	0.016 <i>0.290</i>	
	Yule-Walker (no bridging)	0.000 0.234	0.003 0.263	0.010 0.310	0.024 0.384	0.038 0.455	0.302 2.225	0.075 <i>0.727</i>	0.003 0.250	0.005 0.266	0.009 0.301	0.015 0.340	0.033 0.408	0.013 <i>0.313</i>	
	Yule-Walker (bridging)	0.000 0.234	0.216 1.539	0.445 3.038	3.949 8.001	1.763 6.636	2.996 5.270	1.874 <i>4.897</i>	0.186 1.456	0.190 1.601	0.281 2.071	0.546 3.326	1.331 5.352	0.507 <i>2.761</i>	
	Lomb-Scargle	0.000 0.249	0.185 0.247	0.351 <b>0.247</b>	0.476 0.252	0.529 0.253	0.652 <b>0.251</b>	0.439 <i>0.250</i>	0.096 0.258	0.135 0.261	0.220 0.260	0.307 0.271	0.339 0.284	0.220 <i>0.267</i>	
	Burg for segments	0.000 0.233	0.002 0.248	0.006 0.269	0.012 0.300	0.015 0.331	0.030 0.411	0.013 <i>0.312</i>	0.003 0.245	0.004 0.258	0.007 0.274	0.012 0.302	0.020 0.335	0.009 <i>0.283</i>	

Table 2: Values averaged across all surrogates. For each method and band (LF, HF and LF/HF ratio) two rows are included: the upper one corresponds to  $\overline{NBE}$  whereas the lower row contains  $\overline{CV}$  values. Best values of these magnitudes for each gap pattern (and in average) are emphasized.

results are highly biased when numbers of gaps increases, especially in the HF band due to the growing number of discontinuities. As expected, this effect is more noticeable with 3 s gaps, due to the bigger number of discontinuities.

Linear interpolation substitutes missing samples of the signal with straight segments. There is a general trend to reduced values in power, as can be observed in Figs. 4 and 5. The effect of gaps may be quite erratic, depending on where they fall. It is thus not surprising that the standard deviation of the LF/HF ratio of the estimates increases in presence of more gaps.

#### 4.2. Correlogram

When comparing correlogram with and without bridging the gaps, the latter lacks stability, especially when the percentage of data lost increases. This effect is more noticeable with 3 s gaps, due to the greater number of gaps for a given percentage of missing data. This may be expected, since with many short gaps, the number of contiguous samples between gaps decreases, leaving little data to be used in spectral estimation. Bridging over the gaps uses more of the data.

#### 4.3. Yule-Walker method

For the Yule-Walker method, we used an AR order of 15. Akaike Information Criterion (AIC) suggested model orders greater than 10. Table 2 shows that in general, bridging worsens the results obtained with this method. It should be noted that this result differs from that for the correlogram, where bridging the gaps was beneficial. A likely reason for this will be discussed later.

In Figs. 4 and 5, we only included results obtained with Yule-Walker not bridging over the gaps, and even so, values of standard deviation for 3 s gaps covering the 50% of the surrogates were not included in Fig. 4 to avoid scale problems in the plots. In this case the signal is very fragmented and the estimation gives erratic results.

#### 4.4. Lomb-Scargle method

This algorithm was clearly the one that yielded worst results in terms of bias error when gaps were introduced in the data. The  $\overline{|NBE|}$  values for LF and HF bands were highest, suggesting that it was most affected by gaps. In addition, this method also provided biased estimates when no gaps were present (0%, Figs. 4 and 5), compared to all other methods which all gave similar results at this point. This method was originally devised to detect strong sinusoidal components in noisy and non-uniformly sampled signals, and it is very limited when trying to recover the general shape of a more broadband spectrum.

#### 4.5. Burg algorithm for segments

Burg for segments (AR order 15) was the method that gave the best results, both in terms of mean and standard deviation of power, as shown in Figs. 4 and 5, and of  $\overline{|NBE|}$  and  $\overline{CV}$ , as shown in Table 2. Note that in this table, Burg for segments concentrates most of the best values and that, where it does not, its values are at a short distance from the best. Its ability to combine the

information contained in the remaining segments of data without introducing distortion, gives the most stable results in all circumstances.

In Figs. 4 and 5, the line corresponding to Burg for segments is almost horizontal and always very close to the baseline when plotting mean of power, whereas standard deviation values keep within acceptable values. In Table 2,  $\overline{NBE}$  is under 0.01 and  $\overline{CV}$  is under 0.3 in most cases, both for 3 s and 10 s gaps in LF and HF bands.

## 5. Discussion

In the current study we showed that with an appropriate spectral estimator, gaps in the data need not lead to poor estimates of power in the standard HRV frequency bands, even when up to 50% of the data is missing. Gaps were set in artificially-generated stationary HRV signals. This simulates what one might expect in the presence of sporadic noise, but may not faithfully emulate other important scenarios, such as muscle artefact associated with movement that also leads to a transient change in heart rate, or arrhythmias that have a small but extended impact on sinus rhythm, even after the episode has passed. Our simulations aim to assess the impact in a controlled scenario where the effects of signal processing methods can be observed with little confounding from physiological factors.

It can be seen that for the same percentage of missing samples, 10 s gaps are in general less problematic than 3 s gaps. This is not surprising: for the same percentage of missing data but with shorter gaps, the number of gaps increases. For concatenation or linear interpolation, this means that there are more transients that may cause artefacts. In addition, in algorithms that only work on contiguous data between gaps (correlogram or Yule-Walker without bridging gaps), an increase in the number of gaps leads to a greater loss of useful data which is expected to lead to worse results.

As concatenation is evidently not recommended, it has been proposed to use some form of interpolation to replace the aberrant data [12]. With small amounts of missing data (10 or 20% - which is still more than is usually deemed acceptable), gaps do not necessarily noticeably degrade results for 3 s gaps. There are other forms of interpolation which led to smoother transitions (e.g. cubic splines), but, as said before, resulted in huge excursions in the added segments. We focused instead on methods that exploit the data available, rather than try to impute the missing samples.

Computational load of the methods that try to impute the missing samples depend on the spectral estimation procedure they use. Among the methods that use only the available data, our tests showed that execution times for Lomb-Scargle are approximately one order of magnitude greater than for the correlogram, Yule-Walker or Burg for segments methods. These three are quite similar, with the particularity that in Burg for segments the computational load grows linearly with the percentage of missing samples, while execution times of correlogram and Yule-Walker remain approximately constant.



Lomb-Scargle was specifically devised for detection of sinusoids in non-evenly sampled signal. As it does not require a continuous regular stream of samples, it has been proposed as a good method to deal with gapped data [10]. However, this claim was tested using data where gaps covered up to 15% of the signals, and our experiments showed that performance degrades for greater percentages of missing data, especially in the HF band. The reason for this behaviour is that Lomb-Scargle is based on applying a discrete Fourier transform to the signal, which may include gaps. The Lomb Scargle periodogram uses the relatively high sampling frequency in parts of the signal to compensate for the gaps in others. Gaps thus lead to a reduced average sampling frequency and aliasing effects may occur when the frequency content exceeds half the average sampling rate. This is in contrast to methods such as the correlogram or Burg, where gaps are equivalent to shorter recordings, rather than recordings obtained at a reduced average sampling rate.

The other parametric spectral estimation techniques used in the current work (correlogram, Yule-Walker and Burg for segments), on the other hand, only assume stationarity, and do not aim to describe the signal itself in the frequency domain. They use the available samples to estimate statistical descriptors (e.g. autocorrelation, autoregressive coefficients), and the sampling theorem only needs to be obeyed in the segments that provide samples (assuming regular sampling in these segments). They thus allow for larger gaps than Lomb-Scargle, without the risk of aliasing. The results suggest that this approach is statistically more robust at least in the current application.

Correlogram and Yule-Walker without bridging the gaps treat data segments as entirely disconnected, whether the gap is small (even just one sample) or large. This is associated with some loss of information, as correlation of samples from before and after the gap is ignored. For the specific case of HRV, there may however be an advantage in doing so, as any phase shift in RR intervals after a gap is neglected. As in most estimation methods, there is a balance to be struck between losing useful information (physiologically significant correlation between beats from before and after the gap) and removing noise (distortion from correlations that are not of physiological importance). In the current work with artificially introduced gaps in otherwise (approximately) stationary signals, correlation across the gaps may hold valid information, but even so it is not always helpful in improving spectral estimates.

Comparing the algorithms in different bands, it can be observed that the influence of the distortion introduced by the gaps is more pronounced in the HF band. In general, the results obtained in this band are worse than the ones obtained in the LF band, especially when dealing with short gaps.

It is interesting to note that for the correlogram, bridging the gaps led to better results than without bridging, but for Yule-Walker the opposite was observed. An explanation for this can be found considering the balance between different sources of error. For the correlogram, the autocorrelation function with lags up to 120 samples was estimated. Without bridging the gaps and with many segments of missing data, very few samples would be available to estimate the autocorrelation values at large lags (for 120 lags only

segments without gaps of 30 s duration can be used) and bridging the gaps allows this to be overcome somewhat. For Yule-Walker, only a few samples of the autocorrelation function are needed, corresponding to the order of the model chosen (15 in the current work). Poor estimates of autocorrelation at large lags are therefore irrelevant. However, at lower lags and with short gaps, bridging the gaps means that at some lags some samples will be excluded (because the lagged sample falls in a gap), and at shorter and longer lags, this sample may be used. This somewhat random inclusion and exclusion of samples will lead to some erratic behaviour of the autocorrelation estimates, which reflects in poor spectral estimates with increased high-frequency power. The current results suggest that this effect is more damaging to the spectral estimates than the slight loss in the usable samples that not bridging the gaps entails.

It might be noted that the Burg for segments method, which showed some of the best performance in this spectral analysis also only uses data between gaps, and disregards any correlation in samples that may exist between before and after the gaps.

For the LF/HF ratio, Burg for segments consistently shows the smallest bias error (Table 2). Lomb Scargle, on the other hand, often shows the smallest random error, but this is less impressive when considered in relation to the very large (and downward) bias error this method shows.

## 6. Conclusion

In tests with simulated data, the effect of missing samples, for varying percentages and lengths, on the spectral estimate of signals was analyzed. As expected, when the amount of good data to analyze decreased, the random estimation error increased, and thus the standard deviation of estimates of power in each frequency band and the coefficient of variation also increased.

It was demonstrated that the Burg algorithm for segments is a reliable method for dealing with interrupted RR records, even when a large amount of data is missing. This method outperformed alternatives that are more widely used, including linear interpolation or the Lomb Scargle periodogram. The results showed that even when only 10% of data is missing, Burg for segments provides more robust results than linear interpolation.

Artefacts can have a strong influence on HRV spectra and need to be excluded from HRV analysis. Our results show that even with many gaps (up to 50% of the samples in our study), the recordings need not to be discarded and spectral parameter estimates can still be used. As it was demonstrated by our tests, algorithms such as Burg for segments can provide an accurate estimate from the remaining high quality segments.

While the spectral estimators are assessed here in the context of HRV analysis, the conclusions are probably also valid for many other applications of spectral estimation, especially in the biomedical context, where gaps are often unavoidable. Examples of these are electromyography (EMG) or

electroencephalography (EEG) signals, where gaps can be caused by physiological events (seizures, changes in physiological states, etc.) or noise (patients' movements, muscle contractions, etc.)

### Acknowledgements

This work was supported by the EPSRC [EP/K036157/1] and by a 2018 Salvador de Madariaga grant (*Programa Estatal de Promoción del Talento y su Empleabilidad en I+D+i, Subprograma Estatal de Movilidad, del Plan Estatal de I+D+I*) from the Spanish Ministry of Education, Culture and Sport (MECD).

### References

- [1] E. J. Benjamin, P. Muntner, A. Alonso et al., Heart Disease and Stroke Statistics — 2019 Update: A Report From the American Heart Association, *Circulation* 139 (2019) e56–e66. doi:10.1161/CIR.0000000000000659.
- [2] E. Hoefman, P. J. E. Bindels, H. C. P. M. van Weert, Efficacy of diagnostic tools for detecting cardiac arrhythmias: systematic literature search, *Neth. Heart J.* 18 (11) (2010) 543–551. doi:10.1007/s12471-010-0831-0.
- [3] J. F. Thayer, S. S. Yamamoto, J. F. Brosschot, The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors, *Int. J. Cardiol.* 141 (2) (2010) 122–131. doi:10.1016/j.ijcard.2009.09.543.
- [4] A. J. Camm, M. Malik, J. Bigger et al., Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, *Circulation* 93 (5) (1996) 1043–1065. doi:10.1161/01.CIR.93.5.1043.
- [5] J. Rottenberg, Cardiac vagal control in depression: a critical analysis, *Biol. Psychol.*, 74 (2) (2007) 200–211. doi:10.1016/j.biopsycho.2005.08.010.
- [6] G. G. Berntson, J. Thomas Bigger, D. L. Eckberg, et al., Heart rate variability: Origins, methods, and interpretive caveats, *Psychophysiology* 34 (6) (1997) 623–648. doi:10.1111/j.1469-8986.1997.tb02140.x.
- [7] K. Solem, P. Laguna, L. Sornmo, An efficient method for handling ectopic beats using the heart timing signal, *IEEE Trans. Bio-Med. Eng.* 53 (1) (2006) 13–20. doi:10.1109/TBME.2005.859780.
- [8] G. G. Berntson, J. R. Stowell, ECG artifacts and heart period variability: don't miss a beat!, *Psychophysiology* 35 (1) (1998) 127–132. doi:10.1111/1469-8986.3510127.

- [9] N. Lippman, K. M. Stein, B. B. Lerman, Comparison of methods for removal of ectopy in measurement of heart rate variability, *Am. J. Physiol.-Heart C.* 267 (1) (1994) H411–H418. doi:10.1152/ajpheart.1994.267.1.H411.
- [10] G. D. Clifford, L. Tarassenko, Quantifying errors in spectral estimates of HRV due to beat replacement and resampling, *IEEE Trans. Bio-Med. Electron.* 52 (4) (2005) 630–638. doi:10.1109/TBME.2005.844028.
- [11] R. A. Thuraisingham, Preprocessing RR interval time series for heart rate variability analysis and estimates of standard deviation of RR intervals, *Comput. Meth. Prog. Bio.* 83 (1) (2016) 78–82. doi:10.1016/j.cmpb.2006.05.002.
- [12] M. Peltola, Role of editing of R-R intervals in the analysis of heart rate variability, *Front. Physiol.* 3 (2012) 148. doi:10.3389/fphys.2012.00148.
- [13] S. De Waele, P. M. Broersen, The Burg algorithm for segments, *IEEE Trans. Signal Process.* 48 (10) (2000) 2876–2880. doi:10.1109/78.869039.
- [14] N. R. Lomb, Least-squares frequency analysis of unequally spaced data, *Astrophys. Space Sci.* 39 (2) (1976) 447–462. doi:10.1007/BF00648343.
- [15] M. A. Salo, H. V. Hulkuri, T. Seppanen, Ectopic Beats in Heart Rate Variability Analysis: Effects of Editing on Time and Frequency Domain Measures, *Ann. Noninvas. Electro.* 6 (1) (2006) 5–17. doi:10.1111/j.1542-474X.2001.tb00080.x.
- [16] S. L. Marple Jr, *Digital spectral analysis with applications*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1987.
- [17] J. Burg, Maximum entropy spectral analysis, in: *Proc. 37th Meet. Soc. Exploratorial Geophys.*, Oklahoma City, OK, 1967, reprinted in *Modern Spectrum Analysis*, IEEE Press, 1978.
- [18] J. D. Scargle, Studies in astronomical time series analysis. II - Statistical aspects of spectral analysis of unevenly spaced data, *Astrophys. J.* 263 (1982) 835–853. doi:10.1086/160554.
- [19] J. Mateo, P. Laguna, Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal, *IEEE Trans. Biomed. Eng.* 50 (3) (2003) 334–343. doi:10.1109/TBME.2003.808831.
- [20] A. I. Rincon Soler, L. E. Virgilio Silva, R. Fazan Jr., L. O. Murta Jr, The impact of artifact correction methods of RR series on heart rate variability parameters, *J. Appl. Physiol.* 124 (3) (2018) 646–652. doi:10.1152/jappphysiol.00927.2016.
- [21] A. Choi, H. Shin, Quantitative Analysis of the Effect of an Ectopic Beat on the Heart Rate Variability in the Resting Condition, *Front. Physiol.* 9 (2018) 922. doi:10.3389/fphys.2018.00922.

- [22] K. K. Kim, J. S. Kim, Y. G. Lim, K. S. Park The effect of missing RR-interval data on heart rate variability analysis in the frequency domain, *Physiol. Meas.* 30 (10) (2009) 1039–1050. doi:10.1088/0967-3334/30/10/005.
- [23] P. M. Broersen, S. De Waele, R. Bos, Application of autoregressive spectral analysis to missing data problems, *IEEE Trans. Instrum. Meas.* 53 (4) (2004) 981–986. doi:10.1109/TIM.2004.830597.
- [24] A. Beex, M. Rahman, On averaging Burg spectral estimators for segments, *IEEE Trans. Acoust., Speech, Signal Process.* 34 (6) (1986) 1473–1484. doi:10.1109/TASSP.1986.1164987.
- [25] A. L. Goldberger, L. A. Amaral, L. Glass, et al., Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals, *Circulation* 101 (23) (2000) e215–e220. doi:10.1161/01.CIR.101.23.e215.
- [26] Spectrum: spectral analysis in python, <https://github.com/cokelaer/spectrum>, (accessed 10-aug-2017).
- [27] The wfdb software package, <http://www.physionet.org/physiotools/wfdb.shtml>, (accessed 10-aug-2017).
- [28] G. B. Moody, Spectral analysis of heart rate without resampling, in: *Proc. Computers in Cardiology*, London, 1993, pp. 715–718. doi:10.1109/CIC.1993.378302.
- [29] H. Akaike, A new look at the statistical model identification, *IEEE T. Automat. Contr.* 19 (6) (1974), 716–723. doi:10.1109/TAC.1974.1100705.
- [30] A. Boardman, F. S. Schindwein, A. P. Rocha, A. Leite A study on the optimum order of autoregressive models for heart rate variability, *Physiol. Meas.* 23 (2) (2002), 325–336. doi:10.1088/0967-3334/23/2/308.